The manuscript titled “An increase in the spatial extent of European floods over the last 70 years” by Fang et al., aims to identify large spatio-temporally connected flood events over the last 70 years in Europe using the mHM hydrological model.

Overall, the manuscript is well organized and written. For some of the methods used in the study a better justification should be presented so that the reader is better aware of the reasons for certain choices.

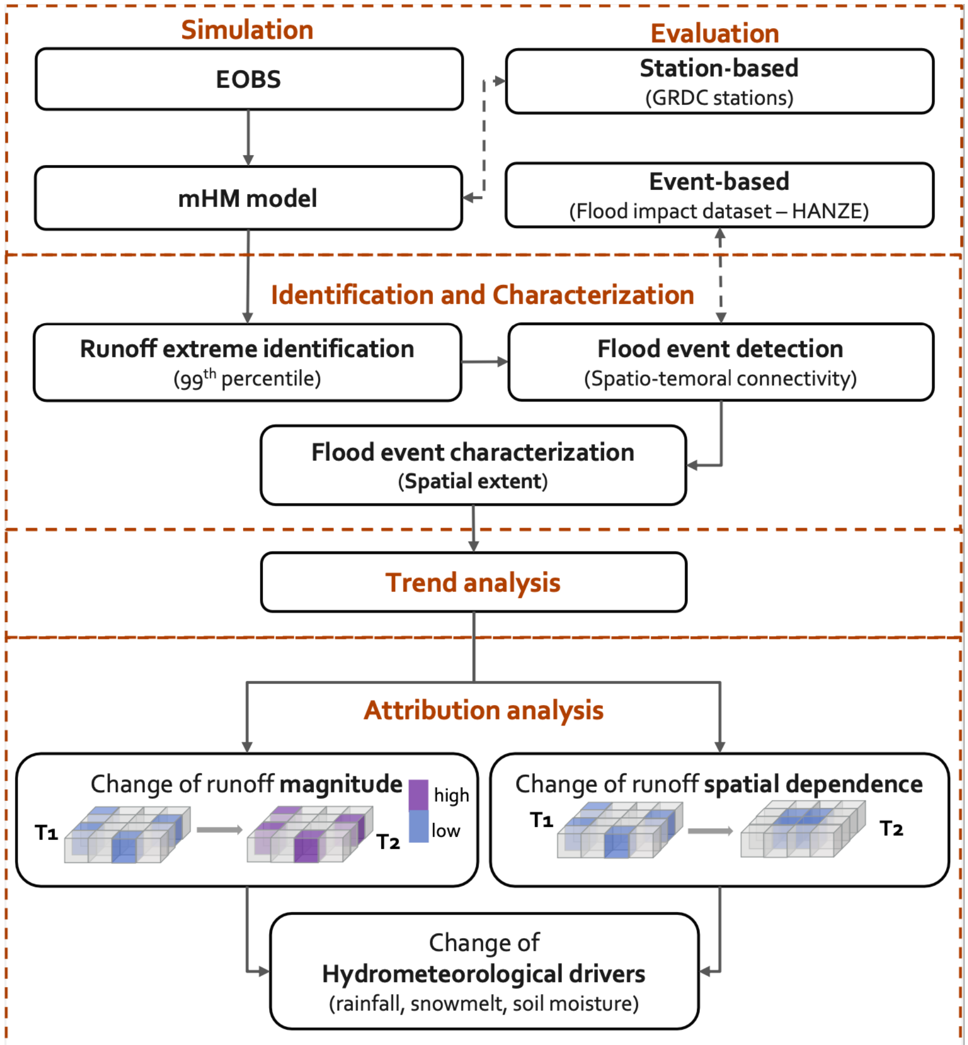
For further details find my comments below.

**General Comments:**

Naturally, as with such a complex study, a lot of data sets and models have been used so that at some stage the reader might get lost what has actually been done. I think it would be helpful to add a schematic/graphical representation showing the inputs, models used and outputs and their ways of comparison in one to two panels to make things clearer.

**Response:**

Thanks for your suggestion and we will add a schematic figure (as shown below) in the manuscript.




Additionally, it is unclear why the period beginning with 1951 was used. The mHM model has a spin-up period of 1940-1959. Hence, I think the analysis should be conducted only after… This would also resolve the issue that as described in L 126 10 years of data are discarded (1981-1990). So the "attribution" could then be done using the two periods 1961-1990 and 1991-2020.

**Response:**

Thanks, it seems the spin-up of mHM was not explained in sufficient detail. The model's initial conditions were created in two decades, namely 1950-1959, and then reading the restart file of 1959 as an initial condition in 1940 run until 1949 (here, using pre1950 EOBS version, see https://surfobs.climate.copernicus.eu/dataaccess/access_eobs.php was used). This means that in total, 20 years were used to create steady-state conditions starting already in 1950. So, the decade

1950-1959 is not affected by the spin up anymore. We use the two time periods as in the manuscript to maximize the trend signal and make use of the full time period available.

Also, the section 2.5 on "attribution" of changes in flood events needs to be re-written. A lot of assumptions are being made in this section, but it remains illusive why certain choices are being made (suggest adding some explanations instead of just citing references that have done the same previously). Additionally, the authors should aim to add some "physical reasoning" to this mainly statistically based attribution exercise. Additionally, the attribution reminds of the use of conditional probabilities. Please elaborate/discuss why this approach has not been considered instead, i.e. what are the advantages of the current approach…

**Response:**

We agree that this section requires a more detailed explanation.

In general, the occurrence of a compound event is shaped by the multivariate probability density function (pdf) associated with the variables describing the event (François & Vrac, 2023; Singh et al., 2021; Sklar, 1973; Zscheischler & Seneviratne, 2017). According to copula theory, a multivariate pdf can be decomposed in the marginal pdfs (i.e. the pdf of the individual variables associated with the compound event) and the copula describing the dependencies among the individual variables. Therefore, the compound event changes are shaped by changes in the marginal pdfs and the dependencies (Bevacqua et al., 2021; Zscheischler & Seneviratne, 2017). We will add this part to section 2.5.

In our specific case, the marginal distributions describe the runoff at individual locations, while the dependency describes the dependency between the runoff at different locations. Accordingly, quantifying the contribution from these two sources is a common practice in compound event research as it provides insights into the origins of the changes. For instance, Bevacqua et al., (2021) applied this approach to differentiate the contribution from rainfall magnitude and rainfall spatial dependence to the extent of winter extreme rainfall. Here, we employ a similar approach to test the hypothesis that flood extent can potentially be influenced not only by changes in runoff magnitude but also by changes in the spatial dependence of high runoff events, as outlined in the introduction (Lines 54-56). We will further clarify the aspects above in the paper.

In the physical reasoning part, utilizing the contribution decomposition method mentioned above, we find that the expansion of flood extent across Northern Europe is primarily driven by the amplified runoff magnitude resulting from increased rainfall and snowmelt (Figure 9). Conversely, the increase of flood extent across central Europe is dominated by the enhanced spatial dependence of runoff extremes, attributed to more widespread heavy rainfall, as outlined in section 3.4. Thus, the decomposition method significantly aids in elucidating the underlying mechanisms behind changes in flood extent.

We agree that conditional probability is commonly used in attribution analysis. For instance, one can investigate the change of flood frequency or magnitude conditioned on different weather types for climate modes. However, in our case the approach would not help much in addressing our main research objective, namely to understand trends in the flood extent.

Finally, I am missing a discussion of the presence of human effects of the hydrology/floods through dams etc. particularly with regard to the conclusions drawn with regard to the attribution of changes in flood extent. With more and large dams that have been built in the recent decades in south of Europe a lot of water will be held back and is no longer available for flooding.

**Response:**

We did not include dams and reservoirs in the model setup, but we agree with the reviewer that it might indeed be a limiting factor, and we will acknowledge this in the revised manuscript. First, we are unaware of a database that would harmonise such information across the entire European domain over time. Additionally, reservoir modelling needs more crucial bathymetric data (information about the underwater topography), accounting for sedimentation processes, which is not available. These data are essential for connection among water elevation, surface area, and reservoir volume, which is necessary for constructing an accurate reservoir model at a large scale, including operational rules. We will newly acknowledge this limiting factor in the paper's discussion, as in reality, the available dam's/reservoir's storage for hydropower generation might influence the seasonal flow patterns and attenuate flood peaks and their impacts, and we propose this topic for future studies.

How do the authors reconcile the reality with the modelled results… Please elaborate and discuss.

**Response:**

The model results have been validated against various independent data sources. As illustrated in section 3.1, we not only compare the mHM simulated runoff with GRDC stations but also validate our identified flood events against an impact-based flood database (HANZE), showing a relatively high consistency. Hence, although our results focus on runoff extremes that may not necessarily lead to impactful floods, they still provide valuable insights applicable to reality.

I also in general agree with most of the comments of Reviewer 2, therefore I will not list similar comments/concerns in this review.

**Specific Comments:**

L 28: Suggest replacing "a future climate" with "future" as a future climate is not the only important factor.

**Response:**

Thanks for your suggestion, we will modify this in the revised manuscript.

L 66: Please add which version of E-OBS was used.

**Response:**

Thanks for your suggestion, the E-OBS version is eobs_v25e and we will add this information to the revised manuscript.

L 70-71: Please specify what percentage "most" refers to and what quantifies as "low" station density. Suggest also showing in the appendix area and stations that have been excluded from the analysis.

**Response:**

Thank you for your suggestion, we agree that this sentence lacks precision. Hence, we revise this statement into: A spatial mask is further applied to exclude catchments with headwater/contributing areas outside the Europe/E-OBS domain (Lehner et al., 2008).

L 82: Please specify what method was used to downscale the resolutions.

**Response:**

Apologies for the confusion. We aggregated (not downscaled) the population data by summing up the population at a 2.5 arc-minute resolution within each 0.125° grid cell. We will clarify this in the revised manuscript.

L 105: Please elaborate with a short sentence, why 0.4.

**Response:**

We have conducted sensitivity tests with the overlap ratio ranging from 0.3 to 0.5 and found that our results are not sensitive to the value of the overlap ratio.

L 112: Please elaborate why 1000 km2.

**Response:**

This choice is motivated by the fact that the mHM resolution is 0.125° × 0.125°, approximately corresponding to an area of 100 $km^2$. We aim to ensure that each large river grid cell has at least 10 upstream cells to facilitate the river routing procedure. Moreover, since our focus is on the climate effect on flood generation, we prioritize grid cells with larger catchment areas. In the future, if mHM has a finer resolution, we may also consider focusing on river grids with lower contributing catchment areas.

L 122: Please elaborate why 0.7.

**Response:**

This ratio was taken from Tarasova et al., (2023), who also applied this ratio to identify the snowmelt-driven floods. We have conducted sensitivity tests with this ratio ranging from 0.6 to 1, and found that our results are not sensitive to the value of this ratio.

L142: suggest replacing “well” with “satisfactory”, as this is the terminology mainly used associated to the use of NSE values.

**Response:**

Thanks for your suggestion, we will revise this in the new manuscript.

L360: I´m not sure if one can conclude that "floods are more widespread in low-lying regions, such as parts than in high mountainous regions like the Alps." As the authors have a priori excluded smaller catchments which would naturally be found in the mountainous areas…

**Response:**

As illustrated in Line 110-112, we intended to exclude non-riverine floods, but not necessarily small-catchment floods. As depicted in Figure 4, the number of flood events in the high-mountainous region is slightly lower; nevertheless, there are still a number of events to support the conclusion.

Figures: In some Figures red and green colour coding is used. Please avoid using this as it is difficult for colour blind readers to discern the Figures. Suggest using "colour-blind safe" colours instead.

**Response:**

Thanks for your suggestion. We will modify the corresponding figures with "colour-blind safe" colors in the manuscript.

**Reference:**


Berghuijs, W. R., Allen, S. T., Harrigan, S., & Kirchner, J. W. (2019). Growing Spatial Scales of Synchronous River Flooding in Europe. *Geophysical Research Letters*, *46*(3), 1423–1428. https://doi.org/10.1029/2018GL081883

Bevacqua, E., Shepherd, T. G., Watson, P. A. G., Sparrow, S., Wallom, D., & Mitchell, D. (2021). Larger Spatial Footprint of Wintertime Total Precipitation Extremes in a Warmer Climate. *Geophysical Research Letters*, *48*(8), e2020GL091990. https://doi.org/10.1029/2020GL091990

François, B., & Vrac, M. (2023). Time of emergence of compound events: Contribution of univariate and dependence properties. *Natural Hazards and Earth System Sciences*, *23*(1), 21–44. https://doi.org/10.5194/nhess-23-21-2023

Jiang, S., Bevacqua, E., & Zscheischler, J. (2022). River flooding mechanisms and their changes in Europe revealed by explainable machine learning. *Hydrology and Earth System Sciences*, *26*(24), 6339–6359. https://doi.org/10.5194/hess-26-6339-2022

Lehner, B., Verdin, K., & Jarvis, A. (2008). New Global Hydrography Derived From Spaceborne Elevation Data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Singh, H., Najafi, M. R., & Cannon, A. J. (2021). Characterizing non-stationary compound extreme events in a changing climate based on large-ensemble climate simulations. *Climate Dynamics*, *56*(5), 1389–1405. https://doi.org/10.1007/s00382-020-05538-2

Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, 9(6), 449-460.

Tarasova, L., Lun, D., Merz, R., Blöschl, G., Basso, S., Bertola, M., Miniussi, A., Rakovec, O., Samaniego, L., Thober, S., & Kumar, R. (2023). Shifts in flood generation processes exacerbate regional flood anomalies in Europe. *Communications Earth & Environment*, *4*(1), Article 1. https://doi.org/10.1038/s43247-023-00714-8

Zscheischler, J., & Seneviratne, S. I. (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, *3*(6), e1700263. https://doi.org/10.1126/sciadv.1700263